14.0 DIETSYS EDIT CHECKING FEATURE

TABLE OF CONTENTS

SECTION 14.0

01/06/94

14.1 **Introduction to the DIETSYS Edit Checking Feature.** A critical step in the analysis process is verifying the integrity of the questionnaire data. The edit checking feature of DIETSYS can assist in this task. This system identifies outliers and invalid codings in the questionnaire data. Unusual or invalid data detected by the edit checking feature can then be corrected prior to analyzing the questionnaire data.

The primary purpose of the DIETSYS Edit Checking utility is to identify errors and outliers in the responses to the diet questionnaire. This then gives the investigator the opportunity to correct or exclude offending records. While error-checking is obviously a component of investigators' data-management, this program provides a standardized mechanism so that one can be sure that the same criteria are being applied to all questionnaires. It does not, of course, preclude investigators from conducting additional scrutiny of their data. Indeed it is hoped that this will assist them in doing so.

The DIETSYS Edit Checking utility is not a substitute for the double-keying method for quality control of data input. DIETSYS will only point out key-punching errors if the error has resulted in an invalid coding of a questionnaire response or a questionably high food frequency. Other keying mistakes could result in erroneous but reasonable coding of the questionnaire responses.

14.2 **Edit Checking Questionnaire Data.** Producing the DIETSYS Edit Report is a simple five step process.

STEP 1: Select **Edit Check a File of Questionnaires** from the DIETSYS main menu.

STEP 2: Instruct DIETSYS to save the DIETSYS Edit Report in a file or send it directly to your printer. The DIETSYS Edit Report is described in detail in Section 14.6.

STEP 3: Enter the names of the files required by the DIETSYS Edit Checking utility. Each of these files is described in Section 14.3.

STEP 4: (Optional) Select **Modify Error/Warning Levels** for Edit-Checks. Section 14.4 describes each variable used for the Error/Warning messages.

STEP 5: Select **Do Edit-Checks** from the "Edit Checking" menu. You will be prompted to set the "Options for Edit-Checking Questionnaires". These options are described in Section 14.5.

14.3 **Files Needed For Edit Checking a File of Questionnaires.** DIETSYS uses the screen shown below to prompt for the files required to edit-check questionnaire data. Follow Steps 1 and 2 in Section 14.2 to access this screen.

```
┌─────────────────────────────────────────────────────────────────┐
│        Files Needed For Edit Checking A File of Questionnaires    │
├─────────────────────────────────────────────────────────────────┤
│ Questionnaire Data File:                                          │
│   C:\HHHQ\DEMO.DAT                                                │
│                                                                   │
│ Edit Report File (*.EDT):                                         │
│   C:\HHHQ\DEMO.RPT                                                │
│                                                                   │
│ Error Summary File (*.ERR):                                       │
│   C:\HHHQ\DEMO.ERR                                                │
│                                                                   │
│ Portion Size Data File:                                           │
│   C:\HHHQ\DIETPORT.V30                                            │
│                                                                   │
│ Questionnaire Configuration File (*.CFG):                         │
│   C:\HHHQ\BRIEF87.CFG                                             │
│                                                                   │
│ Analysis Options File (*.OPT):                                    │
│   C:\HHHQ\BRIEF87.OPT                                             │
│                                                                   │
└─────────────────────────────────────────────────────────────────┘
```

Default file names will be available for some prompts. Other prompts will only contain a default path (a path is the drive and directory, "C:\HHHQ\" is a path).

You must enter file names for those prompts which only contain a path. If necessary, you may enter new paths and/or file names for all prompts. If a path is not specified, the path of the current working directory will be used. To make changes, simply type over the names shown. The LEFT/RIGHT arrow keys and the END key can be used to pass over text which you do not wish to change. As with all file screens in DIETSYS, you may use the UP/DOWN arrows or the ENTER key to accept default file names shown.

A. **Questionnaire Data File.** The coded responses from the questionnaire are stored in this file. This includes the food frequencies and other types of questionnaire data. The DIETSYS Edit Checking utility will generate a report of the coding errors and outliers found in the responses stored in this file.

B. **Edit Report File.** The listing of warnings and errors will be written to this file. DIETSYS will not prompt for the name of the Edit Report File if you instructed DIETSYS to send the listing directly to the printer (Step 2 of Section 14.2).

C. **Error Summary File.** This file is created by DIETSYS while Edit-Checking the questionnaires. Error variables calculated for each questionnaire are stored in this file. Descriptions and file locations for the variables in the Error Summary File are contained in Section 14.18.

D. **Questionnaire Configuration File.** DIETSYS uses the information stored in the CFG File to determine the identity and sequence of all questions on your questionnaire (Section 11.1). Enter the name of the CFG File which corresponds to the questionnaires being edit-checked.

E.   **Analysis Options File.** The Analysis Options (Section 16, Appendix A) and Group Information (Section 15.14) are stored in this file. Group Information contains data for both the analysis and edit-checking of the defined food groups. Be sure to enter the name of the options file which contains the correct Group Information data so that the DIETSYS Edit Checking utility may produce the correct reasonableness checks for the group frequencies.

14.4   **Modify Error/Warning Levels.** Some edit-checking errors and warnings are based on patterns of responses rather than an individual response. For example, it is perfectly legitimate for a respondent to select medium for one food portion size, but it suggests negligence if medium is selected for all foods. The "Error Level" for "Coded as Small" is a percentage of all foods asked on the questionnaire. If a respondent answers medium for a percentage of foods greater than the Error Level, an error message will be generated (and an error flag given in the Error Summary File). If the percentage of mediums is less than the Error Level but greater than or equal to the Warning Level, a warning message will be generated (and a warning flag given in the Error Summary File). Naturally, this Error Level must be greater than its Warning Level counterpart.

DIETSYS uses the screen shown below to allow you to set new defaults for the Error/Warning Levels. Follow Steps 1-4 in Section 14.2 to access this screen.

```
┌────────────────────────────────────────────────────────────────────────┐
│               ┌────────────────────────────────────┐                     │
│               │        Error/Warning Levels         │                     │
│               └────────────────────────────────────┘                     │
│                                                                          │
│  1. TOO FEW FOODS/DAY (MALES)....  Warning <  6 foods       Error <  5 foods  │
│                                                                          │
│  2. TOO FEW FOODS/DAY (FEMALES)..  Warning <  5 foods       Error <  4 foods  │
│                                                                          │
│  3. TOO MANY FOODS/DAY (ERROR)...  Males   > 30 foods    Females > 30 foods  │
│                                                                          │
│  4. QUESTIONABLY HIGH FREQS......  Warning >  0 foods       Error >  2 foods  │
│                                                                          │
│  5. FOODS SKIPPED...............   Warning =  10 %          Error >  15 %  │
│                                                                          │
│  6. CODED ONCE/TIME UNIT........   Warning =  70 %          Error >  75 %  │
│                                                                          │
│  7. CODED AS SMALL..............   Warning =  85 %          Error >  99 %  │
│                                                                          │
│  8. CODED AS MEDIUM ............   Warning =  94 %          Error >  99 %  │
│                                                                          │
│  9. CODED AS LARGE..............   Warning =  85 %          Error >  99 %  │
│                                                                          │
└────────────────────────────────────────────────────────────────────────┘
```

Note that number six will read "CODED USING THE SAME FREQUENCY" if the questionnaire is in the categorical format. The levels shown in the sample screen above are the default values distributed with the DIETSYS software. Each type of error shown is discussed in Section 14.6.

Use the UP/DOWN arrow keys to move within a column. You may use the TAB key to move from the left column to the right, and SHIFT-TAB to move the cursor from the right to the left column. If you have made changes you do not wish to keep, press ESC. The levels will be

restored to the values shown when you entered the screen. Press ALTX to exit and save the values displayed.

14.5 **Options for Edit-Checking Questionnaires.** Once you have selected "Do Edit-Checks" from the DIETSYS Edit Checking menu by following steps 1-5 in Section 14.2, DIETSYS will display the following screen.

```
┌──────────────────────────────────────────────┐
│        Options for Edit-Checking Questionnaires │
└──────────────────────────────────────────────┘

        Create a FIX File = OFF

            Edit Report = FULL REPORT
```

A. **Create a FIX File.** The "Fix" File is a copy of the questionnaire file with adjustments made in an attempt to reduce the impact of reported high food frequencies that exceed the Reasonable Limit. Any food that has a high frequency itself or belongs to a group with a high frequency will be adjusted. The adjustment consists of reducing the portion by one serving size (Extra Large to Large, Large to Medium, Medium to Small). No adjustment is made to foods which have a Small portion size.

DIETSYS follows the steps listed below to make the adjustments for the Fixed data. See Section 15.14 for information regarding Food Groups, Group Reasonable Limits and the Fix option for Food Groups. See Sections 12.10 and 12.27 for information regarding Reasonable Limits for individual food items.

1. Food Group Frequencies are checked first. A Food Group Frequency is the sum of the frequencies of all members of the group (Section 15.14). Each group frequency is compared to the group's Reasonable Limit. Only Food Groups with non-zero Reasonable Limits are checked since a Reasonable Limit of zero indicates unlimited (that is, no frequency check should be done by DIETSYS).

   If the group frequency is greater than its Reasonable Limit and the Fix parameter for the Group is set to ON (Section 15.15), the serving size of each member food is reduced by one serving size (except smalls).

   The serving size of any one food will only be adjusted once. That is, if a food is a member of two groups which fit the criteria to be "Fixed", that food will only be adjusted once.

2. The portion size of each food not adjusted in Step 1 which has a frequency greater than its Reasonable Limit will be adjusted. That is, the serving size will be reduced by one size (except smalls).

B. **Edit Report.** This option affects the amount of output created by the DIETSYS Edit Checking system. The following settings are available for this option.

1. Summary Only. Outliers and errors are tabulated for all questionnaires in the file but not reported for each individual. See Section 14.17 for more information.

2. Full Report. All outliers and errors discovered for each individual respondent are listed as well as the summary described above. Each type of error and warning message is discussed in Section 14.6.

14.6 **Individual Errors Reported in the DIETSYS Edit Report.** The errors and warning messages discussed in this section are only included in the DIETSYS Edit Report if the Edit Report option is set to Full Report (see Section 14.5). Each questionnaire in the data file will be edit-checked in sequence. If an individual questionnaire contains questionable data, all errors and warnings for that questionnaire will be listed. The Respondent ID will be listed first, followed by the warning and error messages. Questionnaires which do not have questionable data will not be listed in the report.

All possible error/warning messages for individual questionnaires are listed below. The Warning and Error levels can be set by following the instructions in Section 14.4. Detailed descriptions of the criteria used to set the warning/error flags can be found in the sections listed below.

Section 14.7   Too many food questions skipped
Section 14.8   Too many foods eaten daily
Section 14.9   Too few foods eaten daily
Section 14.10   Too many foods coded in a frequency column
Section 14.11   Too many foods coded as once per time unit
Section 14.12   Too many foods coded with the same serving size
Section 14.13   Food coding errors
Section 14.14   Questionably high food frequencies
Section 14.15   Food Groups with questionably high frequencies
Section 14.16   HHHQ Questions with coding errors

14.7 **Too many food questions skipped:**
x out of y food questions asked in the Main Food List.

$x =$ Number of foods skipped. Foods in Open Ended food lists are not considered when counting skipped foods. A food is defined as skipped if frequency for that food is coded as Missing or Error. For Non-Categorical food frequencies, '99' for the Number of Times segment of frequency constitutes a skipped food. See Section 10.3 for the coding of Categorical and Non-Categorical food frequencies.

$y =$ The number of food questions asked in the Main Food List and does not include foods in Open Ended Food sections of the questionnaire.

This message will be labeled an "Error" if the percentage of foods skipped is greater than the Error Level. A "Warning" message will be generated if the number of foods is greater than or equal to the Warning Level but less than the Error Level. The number of foods skipped for an individual will not be reported in the DIETSYS Edit Report if the percentage of skipped

foods is not at least equal to the Warning Level. See Section 14.4 for instructions in setting the Error and Warning Levels.

14.8   **Too many foods eaten daily: *x***

$x =$ The DIETSYS calculation of the number of foods eaten daily. This is the average daily frequency of foods designated as Solid Foods in the Foods Database. See Section 12.9 for more information regarding the Solid Food flag.

There is no Warning Level for the "Too many foods eaten daily" flag. This message will be generated if the number of foods eaten per day is greater than the Error Level for the respondent's gender (Section 14.4).

It is recommended that an error be generated for both men and women if more than 30 foods are eaten per day. However, you may set different levels based on sex (Section 14.4). These levels are based on a 98 item questionnaire; you may wish to consider a different error level for shorter questionnaires.

14.9   **Too few foods eaten daily: *x***

$x =$ The DIETSYS calculation of the number of foods eaten daily. This is the average daily frequency of foods designated as Solid Foods in the Foods Database. See Section 12.9 for more information regarding the Solid Food flag.

This message will be labeled "Warning" if the number of foods eaten daily is less than the Warning Level but not less than the Error Level. This message will be labeled an "Error" if the number of foods is less than the Error Level (Section 14.4).

The default Warning and Error Levels for "Too few foods eaten daily" are based on a 98 item questionnaire and are not recommended for brief, 60 item questionnaires.

14.10 **Too many foods coded in frequency column *x*:**
        y out of *z* Foods Analyzed

$x =$ the number of the food frequency column as it appears on the printed questionnaire. For this edit check, the columns will be considered as 1 through n, from left to right across the page. DIETSYS determines the column number based on the food frequency coding scheme information stored in the CFG File. See Sections 14.18 and 11.12 for more information.

$y =$ the number of foods with a mark in this frequency column

$z =$ the number of foods analyzed. These are foods on the questionnaire which have a valid frequency. This does not include foods which have frequency codes of Missing or Error. Foods from both the Main Food List and Open Ended Food sections are considered.

This message will only be generated for questionnaires in the Categorical format (Section 10.3). This message will be labeled an "Error" if the percentage of foods coded with a single frequency column is greater than the "Coded Using the Same Frequency" Error Level. A "Warning" message will be generated if this percentage is greater than the Warning Level but less than or equal to the Error Level. See Section 14.4 for instructions in setting the Error and Warning Levels.

14.11 **Too many foods coded or imputed as once per time unit:**
x out of y Foods Ever Eaten

This message will only be generated for Questionnaire Data Files in the Non-Categorical format (Section 10.3). This edit-check is designed for three-character food frequencies, that is, the food frequencies of Non-Categorical questionnaires.

$x =$ the number of foods coded as once per day ('011'), once per week ('012'), once per month ('013'), once per year ('014'), as well as '991', '992', '993', and '994'

Food frequencies coded as '991', '992', '993', or '994' have a missing number of times with a valid unit of time. "Once" is imputed for the number of times of these frequency codes.

$y =$ the number of foods ever eaten. These are foods reported on the questionnaire with a valid, non-zero frequency. This count does not include foods which have frequency codes of Missing ('999'), or never. Foods from both the Main Food List and Open Ended Food sections are considered.

This message will be labeled an "Error" if the percentage of foods coded or imputed as "once per" is greater than the Error Level. A "Warning" message will be generated if this percentage of foods is greater than or equal to the Warning Level but less than or equal to the Error Level. See Section 14.4 for instructions in setting the Error and Warning Levels.

14.12 **Too many foods coded with ... serving size:**
x out of y Foods Ever Eaten

The serving size will be indicated (small, medium, large, or x-large).

$x =$ The number of the foods coded with the serving size indicated. If the serving size is "medium", this number will include the number of foods imputed as medium because serving size is coded as Missing or Error.

$y =$ the number of foods ever eaten. These are foods reported on the questionnaire with a valid, non-zero frequency. This count does not include foods which have frequency codes of Missing, Error, or never. Foods from both the Main Food List and Open Ended Food sections are considered.

Note that this message will not be generated if medium is imputed for all foods because the MedOnly analysis option is ON or serving size is not asked. See Section 16A.17 for more information on the MedOnly analysis option.

14.13 **Food Coding Errors**

   xxx Food Name
    Coded Response: "*yyyy*" . . . . Location: (column)

   Open Ended Food *z*
    Coded Response: "*yyyyyy*" . . . . Location: (column)

The first type of Food Coding Error listed is for a food which appears in the Main Food List of the questionnaire. The second message shows what will be printed for foods in Open Ended Food sections of the questionnaire.

   xxx = The Food ID. The Food ID corresponds to the food's database sequence number (Section 12.3).

   yyyy = The coded response for frequency and serving size for this food. For foods in the Main Food List, this will be 1 to 4 characters depending on whether serving size is asked and whether the frequency is in the Categorical or Non-Categorical format (Section 10.3). Two additional characters will be included in the coded response for Open Ended foods. These two characters are the Open Ended Food Identifier (Section 10.6). The coded response is enclosed in quotation marks so that you will be able to see embedded blanks.

   z = Open Ended foods are numbered from 1 to n, where n is the number of Open Ended foods in all sections of the questionnaire. That is, the total number of Open Ended foods coded (with a frequency or as unstated). *z* is this number.

Location refers to the file location of the coded response. Card and column will be listed if the Questionnaire Data File is in Card Format.

14.14 **Questionably High Food Frequencies**

   xxx Food Name
    Coded Response: "*yyyy*"
  Location: (column)

   Open Ended Food *z*
    Coded Response: "*yyyyyy*"   Location: (column)

A "High Food Frequency" is defined as a food which has a weekly frequency greater than the Reasonable Limit for the food. A Reasonable Limit is set for each food in the DIETSYS Foods Database (Sections 12.10 and 12.27)

The first type of food listed is a food from the Main Food List of the questionnaire. The second message shows what will be printed for foods in Open Ended Food sections of the questionnaire.

  xxx = The Food ID. The Food ID corresponds to the food's database sequence number (Section 12.3).

yyyy	=	The coded response for frequency and serving size for this food. For foods in the Main Food List, this will be 1 to 4 characters depending on whether serving size is asked and whether the frequency is in the Categorical or Non-Categorical format (Section 10.3). Two additional characters will be included in the coded response for Open Ended foods. These two characters are the Open Ended Food Identifier (Section 10.6). The coded response is enclosed in quotation marks so that you will be able to see embedded blanks.
z	=	Open Ended foods are numbered from 1 to n, where n is the number of Open Ended foods in all sections of the questionnaire. That is, the total number of Open Ended foods coded (with a frequency or as unstated). *z* is this number.

Location refers to the file location of the coded response. Card and column will be listed if the Questionnaire Data File is in Card Format.

14.15 Food Groups with Questionably High Frequencies
Group Name   x servings per day

Food Groups are defined in the Analysis Option File (Section 15.14). Each Food Group may have a Reasonable Limit (weekly). If the value of the Reasonable Limit is zero, a reasonableness check on the group frequency is not done. Otherwise, the frequencies of all foods in the group are summed. This group frequency is then compared to the group's Reasonable Limit. If the group frequency is higher, the warning message shown above will be listed in the edit report.

Members of groups with questionably high group frequencies will be adjusted according to the "Fix" algorithm if the Fix group option is ON and the "Create a Fix File" edit-checking option is ON. See Section 14.5 for more information on the "Create a Fix File" option. See Section 15.15 for more information on the Fix group option.

14.16 HHHQ Questions with Coding Errors
Name of HHHQ Question
Location: (column)   Code: (coded response)

Coding Errors refer to responses with invalid codes, such as a blank when only a '1', '2', or '9' are expected. DIETSYS will not analyze files containing coding errors. They must be corrected.

Location refers to the file location of the coded response. Card and column will be listed if the Questionnaire Data File is in Card Format. The coded response will be printed within quotation marks and is the data located in the Questionnaire Data File at the specified Location.

Refer to the coding manual for your version of the questionnaire to check the coding rules for each HHHQ Question. If you are using an NCI version of the questionnaire, look up the HHHQ Question by the file location. Codebooks for the NCI Questionnaires may be found in the following sections of this document.

Paper Questionnaire	Codebook
SCAN92	Section 7
FULL87	Section 6
BRIEF87	Section 8

Interactive Interview	Codebook
LONG	Section 6
SHORT	Section 9

If you have designed your own questionnaire, refer to Section 10, Appendix A for coding instructions (the HHHQ Questions are listed by name).

14.17 The summary on the following page is printed at the end of the DIETSYS Edit Checking Report. Each message printed shows a population total of a warning or error. The warning and error types are described in Sections 14.5 through 14.16. The "Questionnaires with questionable data" count at the bottom of the summary indicates the number of different questionnaires with at least one error (questionnaires with only warnings do not count toward this total).

A similar summary page is printed with the DIETSYS Analysis Report. Note that the "foods per day" and "high frequency" warning/error counts in the Analysis Report summary may differ from the Edit Report summary. These discrepancies occur because the food frequencies calculated by the DIETSYS Nutrient Analysis are adjusted according to the algorithms used to implement the Analysis Options (Section 16, Appendix A). The food frequencies used by the Edit-Checking system for the "foods per day" and "high food frequency" flags are simply the reported frequencies multiplied by the food's Seasonality Factor (Section 12.8). The Analysis Options which affect the individual food frequencies include VegAdjust, CerealAdjust, FruitAdjust, and RestAdjust. Refer to Section 16 for a detailed description of each Analysis Option. In addition, if any foods are Dropped from the analysis, the "foods per day" and "high frequency" warning/error counts will be affected (Section 15.13).

DIETSYS Edit Report Summary

Number of Questionnaires Checked: 169
Number with one or more warning or error flags: 34

Error Description	Records With Error Number	Percent
Foods With Invalid Coding:		
Error: 1 or more miscodes.	0	0.0
Invalid Coding for questions Other than Food List:		
Error: 1 or more miscodes.	0	0.0
Foods Skipped: (missing food frequencies):		
Warning: 10% to 15%.	0	0.0
Error: More than 15%	1	0.6
Too few foods per day:		
Males: Warning: less than 6	5	6.0
Error: less than 5	8	9.5
Females: Warning: less than 5	4	4.7
Error: less than 4	2	2.4
Error: Too many foods per day:		
Males: More than 30.	0	0.0
Females: More than 30.	0	0.0
Foods coded using the same frequency:		
Warning: 70% to 75%.	1	0.6
Error: More than 75%	0	0.0
Foods coded as small serving size:		
Warning: 85% to 99%.	1	0.6
Error: More than 99%	0	0.0
Foods coded as medium or no serving size coded:		
Warning: 94% to 99%.	10	5.9
Error: More than 99%	4	2.4
Foods coded as large serving size:		
Warning: 85% to 99%.	0	0.0
Error: More than 99%	0	0.0
Questionably high frequencies:		
Warning: 1 to 2 questionable foods	3	1.8
Error: 3 or more questionable foods. . . .	0	0.0
Warning: 1 or more high group frequencies.	0	0.0
Questionnaires with questionable data	14	8.3

14.18 **Error Summary File.** There are 24 warning and error flags which are calculated for each questionnaire and stored in the Error Summary File. This file is an ASCII text file. One line of data per questionnaire is written to this file.

This file is space delimited. That is, the 24 variables are separated from each other with one blank. All floating point variables are formatted as x.xx. The number of digits on the left of the decimal point will vary according to the magnitude of the value.

The variables appear in the following order:

1. **Respondent ID**

    This is a 10 character field.

2. **Number of Food Frequencies Skipped**

    Integer Format.

    Foods in Open Ended food lists are not considered when counting skipped foods. A food is defined as skipped if frequency for that food is coded as Missing or Error. For Non-Categorical food frequencies, if '99' is coded for the Number of Times segment of frequency the food is counted as skipped. Note that '01' will be imputed for number of times if the time unit is day, week, month or year (that is, '991', '992', '993' or '994'). Foods with an imputed number of times will be counted as skipped but still contribute to the nutrient estimates. See Section 10.3 for more information on Categorical and Non-Categorical food frequencies.

3. **Number of Foods With No Serving Size**

    Integer Format.

    This is the number of foods with a non-zero frequency which have a serving size of Missing or Error. Foods with Missing, Error, zero, or "never" coded for frequency will not be counted in this sum.

    The value of this variable will be zero if serving size was not asked on the questionnaire.

4. **Foods Per Day**

    Floating Point Format.

    This variable is the summed daily frequency of all solid foods. Each food in the DIETSYS Foods Database has a Solid Food indicator. All foods analyzed for a respondent which have 'Y' for the Solid Food flag will be counted toward the Foods Per Day calculation.

5 - 8. Percent of Foods Coded With ... Serving Size

Floating Point Format.

These variables are only relevant for questionnaires which included serving size with the food frequencies. If serving size was not asked, a '0' will be coded in the Summary Error File for each of these variables.

Each percentage is based on the foods with a non-zero frequency. Each is a percentage of the foods with a non-zero frequency which have the particular serving size. The Percent of Foods Coded as Medium includes foods with Missing or Error coded for serving size (medium is imputed for these foods).

The serving size variables appear in the Summary Error File in the following order:

5. Percent of Foods Coded as Small
6. Percent of Foods Coded as Medium
7. Percent of Foods Coded as Large
8. Percent of Foods Coded as Extra Large

9. Number of Foods Coded as Once Per Any Time Unit

Integer Format.

This variable is only relevant for Questionnaire Data Files in the Non-Categorical format (Section 10.3). If the Questionnaire Data File is in Categorical format, zero will be the value of this variable for all respondents in the Summary Error File.

The number of foods coded as once per time unit includes foods with '01' for the number of times as well as foods with a missing number of times ('99') and a valid time unit. That is, this variable is the total number of foods with any of the frequency codes below.

'011' or '991' Once Per Day
'012' or '992' Once Per Week
'013' or '993' Once Per Month
'014' or '994' Once Per Year (unless a period other than year is specified in the CFG File)

10 - 18. Percent of Foods Coded in Frequency Column n

Floating Point Format.

The eleven variables listed below are only relevant for Questionnaire Data Files in the Categorical Format (Section 10.3). '0.00' will be coded in the Summary Error File for each of these variables if the data file is in the Non-Categorical format. If there are fewer than nine frequency columns, '0.00' will be coded for the unused columns.

10. Percent of Foods Coded in Frequency Column 1
11. Percent of Foods Coded in Frequency Column 2
12. Percent of Foods Coded in Frequency Column 3
13. Percent of Foods Coded in Frequency Column 4
14. Percent of Foods Coded in Frequency Column 5
15. Percent of Foods Coded in Frequency Column 6
16. Percent of Foods Coded in Frequency Column 7
17. Percent of Foods Coded in Frequency Column 8
18. Percent of Foods Coded in Frequency Column 9

These percentages are calculated using a denominator equal to the total number of foods minus foods coded as missing or error. These include foods in the Main Food List and foods in Open Ended Food sections.

The "Foods Coded in Frequency Column 1" are the foods which are coded with the first frequency input character defined in the CFG File. It is recommended that the first input character defined be the character used for the left-most column ("Column 1") on the questionnaire. See Section 11.12 for instructions on setting the food frequency coding schemes.

These percentages are based solely on the column position, not on the frequency value attributed to that column. For example, foods on pages 4 through 8 of the SCAN92 questionnaire with a frequency of "2+ per day" are counted as Column 9 frequencies. Foods on page 9 of SCAN92 coded as "6+ per day" are also counted as Column 9 frequencies.

If a respondent indicated a large percentage of responses in the "Never" column of your questionnaire (SCAN92 Frequency Column 1), it may indicate that the food list is inappropriate for his/her diet. A large percentage of responses in any other single column may indicate a careless respondent.

19. Number of Foods with Questionably High Food Frequencies

Integer Format.

Foods with Questionably High Food Frequencies are foods with a reported frequency greater than the Reasonable Limit set for the food in the DIETSYS Foods Database. Section 12.10 contains an explanation of the Reasonable Limit database variable. Section 12.27 describes how to change the value of the Reasonable Limit for any food.

Note that questionably high frequencies of Grouped Foods are not considered here.

20. Number of Questionably High Group Frequencies

Integer Format.

A Food Group Frequency is the sum of the frequencies of all members of the group (Section 15.14). Each group frequency is compared to the group Reasonable Limit (Section 15.15). If the group frequency is higher, it is counted here as questionably high.

21. Fix Option Setting.

Integer Format.

This variable will have a value of one or zero. The value will be the same for all respondents. One indicates that the "Create a Fix File" option was ON during the Edit Checking run which created this Summary Error File. Conversely, zero means this option was OFF. See Section 14.5 for more information regarding this option.

A Fix Option Setting of one and a value of zero in field #22 (the number of fixes performed for this respondent) indicates that no fixes were deemed necessary by the Edit Checking system. Otherwise, a value of zero for field #22 would be ambiguous and might indicate that the fix option was simply not implemented.

22. Number of Fixes Performed for this Respondent

Integer Format.

This is the number of foods which were changed according to the algorithm for creating a Fix File. Each food will only be changed once. See Section 14.5 for more information on creating a Fix File.

23. Total Number of Warnings for this Respondent

Integer Format.

The number of warnings include the following:

1. Too many food questions skipped
2. Too few foods eaten daily
3. Too many foods coded with the same frequency, that is, coded in the same frequency column
4. Too many foods coded as once per time unit
5. Too many foods coded with a small serving size
6. Too many foods coded or imputed with a medium serving size
7. Too many foods coded with a large serving size
8. Too many foods coded with an extra large serving size
9. Too many foods with questionably high food frequencies
10. One or more questionably high group frequencies

24. Total Number of Errors for this Respondent

Integer Format.

The number of errors includes the following:

A.   The number of invalid codes in the coded responses of food frequencies and HHHQ Questions.

B.   The number of reasonableness edit checks which exceed the corresponding Error Level (Section 14.5). These include the following:

    1.   Too many food questions skipped
    2.   Too many foods eaten daily
    3.   Too few foods eaten daily
    4.   Too many foods coded with the same frequency, that is, coded in the same frequency column
    5.   Too many foods coded as once per time unit
    6.   Too many foods coded with a small serving size
    7.   Too many foods coded or imputed with a medium serving size
    8.   Too many foods coded with a large serving size
    9.   Too many foods coded with an extra large serving size
    10. Too many foods with questionably high food frequencies